心理科学进展 2018, Vol. 26, No. 1, 62–71
Advances in Psychological Science DOI: 10.3724/SP.J.1042.2018.00062

·研究前沿(Regular Articles)·

声调的范畴知觉及其神经机制*

吴 倩[1] 王韫佳[1,2]
([1]北京大学中国语言文学系) ([2]北京大学中国语言学研究中心, 北京 100871)

摘 要 语音范畴知觉是指听者能够区分不同音位范畴的刺激而不能区分同一范畴内的不同刺激。声调知觉的范畴化程度可能与刺激的声学相似度有关, 相似度越高则知觉的范畴化程度越低。除了声调本身的特征外, 影响声调知觉范畴化的因素还包括母语背景、年龄、刺激所在语境和刺激属性(语言和非语言)。电生理技术的使用加深了声调范畴知觉的研究, 并有助于解决一些长期存在争议的理论问题。

关键词 声调; 范畴知觉; 神经机制
分类号 B842

语音的范畴型知觉模式最早是由 Liberman 及其同事提出的(Liberman, Harris, Hoffman, & Griffith, 1957), 该术语是指在声学连续统中, 连续的语音刺激被感知为若干离散范畴中的一个, 听者能够区分属于不同范畴的刺激, 但不能很好地区分属于同一范畴的刺激(Studdert-Kennedy & Shankweiler, 1970)。与范畴型知觉相对的是连续型知觉, 即连续的语音刺激不能被感知为离散范畴中的一个, 听者对等间距刺激的区分度始终是相同的。

早期的语音范畴知觉研究主要集中在音段上, 例如, 大量研究表明人们对辅音的感知属于范畴型感知, 对元音的感知则属于连续型感知(例如, Fry, Abramson, Eimas, & Liberman, 1962; Liberman et al, 1957)。1976 年, 王士元发现, 被试对北京话阴平和阳平的感知也属于范畴感知(Wang, 1976), 从此声调的范畴感知问题也成为语音范畴知觉研究的热点之一。进入新世纪之后, 随着脑神经科学的发展, 电生理学的研究方法使得探究声调知觉行为背后的神经机制成为可能。本文将对有关声调范畴化问题及其影响因素的研究进行评述。

1 声调知觉的范畴化研究

1.1 行为研究

由于技术的原因, 早期的范畴感知研究都是基于行为实验的, 对刺激的辨认和区分是范畴感知实验的两个组成部分。典型的范畴型感知在辨认率曲线和区分正确率曲线上的特点分别是: (1)跨范畴边界的两个相邻刺激的辨认率差别远远大于处于边界同侧的两个相邻刺激的辨认率差别; (2)在声学距离相同的情况下, 跨范畴的两个刺激的区分正确率高于处于范畴同侧的两个刺激的区分正确率(Liberman et al., 1957)。

关于声调的范畴知觉, 学界已有大量行为实验的研究。声调的感知究竟是范畴化的还是连续型的, 这一问题在 20 世纪 70 年代后期曾经有过激烈争论。王士元通过研究北京话阴平和阳平的知觉, 率先提出了声调知觉是范畴型感知的假设(Wang, 1976)。然而, Abramson (1979)对泰语声调的研究表明, 泰语中三个平调的辨认虽然是范畴型的, 但区分曲线上都未出现范畴边界处的区分高峰, 因此他认为泰语的平调感知是非范畴型感知。

进入 21 世纪, 声调的范畴感知问题再次引起

收稿日期: 2016-11-29
* 教育部人文社会科学重点研究基地 2013 年重大项目“汉语声调认知的实验研究——声学变异、范畴感知与连读变调” (13JJD740002)和国家社科基金项目“汉语和非汉语母语者加工普通话声调的 ERP 研究” (11YB047)。
通信作者: 王韫佳, E-mail: wangyunjia@pku.edu.cn

许多研究者的关注。在汉语普通话和汉语方言声调的研究中，范畴型感知的实验结果屡见不鲜(例如，蔡雯清，2016; Francis, Ciocca, & Ng, 2003; Hallé, Chang, & Best, 2004; 刘思维, 2015; Peng et al., 2010; 覃夕航, 2012; Shen, 2015; Wang, 1976; 王韫佳，李美京，2010; 王韫佳，覃夕航，2015; Xi, Zhang, Shu, Zhang, & Li, 2010; Xu, Gandour, & Francis, 2006; 张林军，2010)。然而，也有研究者得到了非范畴型的结果，例如 Francis 等人(2003)对粤语三个平调以及两个升调的研究结果都与泰语平调的感知结果类似。王韫佳和覃夕航(2015)对普通话阳平−上声的感知研究也得到了辨认率和区分率不一致的结果，对此她们提出一个假设，即调形相似(声学相似)的声调之间不存在清晰的区分范畴边界。根据该假设，刘思维(2015)在对北京话、重庆话和天津话声调的感知研究中区分了相异调形和相似调形，结果表明，相异调形的感知呈现出典型的范畴化特征，但相似调形的感知结果却呈现出弱范畴化或非范畴化的特征。蔡雯清(2016)对合肥话声调的研究结果也支持王和覃提出的假设。

关于声调知觉是否具有范畴化特征，前人的研究结果出现了不一致。造成前人结果不一致的一个重要原因是前人所用的声调类型不同，Francis 等人(2003)认为，曲折调的感知呈现范畴型特征，而平调的感知则具有连续型特征。然而，粤语两种升调的感知(Francis et al., 2003)以及普通话阳平−上声连续统的感知(王韫佳，覃夕航，2015)却得到了与泰语平调相类似的结果。由此可见，声调知觉的范畴化程度可能并不仅仅与声调的类型(平、升、降以及更复杂的曲拱类型)相关，而是与两个声调的音高曲拱是否具有相似性相关：两个声学相似度低的声调之间可能存在较为清晰的知觉范畴边界，而声学相似度高的声调之间则不存在清晰的知觉边界。因此，相比于以往在声调范畴研究中区分平调和曲折调，区分相似调形和相异调形(这里的调形特指声调在声学层面的音高走向，而不是音系学层面的声调类型)更便于解释知觉结果。

声调连续统的设计也会影响声调的知觉模式。对声调的知觉，实际上是对某种声学线索组合的知觉(比如终点音高、起点音高等)，不同的连续统设计可能会对不同的声学线索进行改变，而被试对不同声学线索的敏感程度是不同的，这就会造成知觉结果的差异。例如，王韫佳和覃夕航(2015)比较了不同的刺激设计对阳平−上声连续统感知的影响，她们以降升调的拐点位置为自变量，在不同的下降段降幅和上升段升幅条件下观察声调知觉的范畴化程度，结果表明，在特定下降段降幅和上升段升幅条件下，阳平和上声的知觉才会呈现弱范畴感知的特点。

另一个值得注意的结果是，在多数针对两个声学层面具有较大相似性的声调的行为实验中，辨认实验的结果都远远好于区分实验的结果。从前人的结果看，区分实验结果在很大程度上与记忆机制有关。首先，记忆中的音高衰减会影响一部分区分结果(Francis & Ciocca, 2003)。其次，人们对于声调在语流中的各种变异的记忆也会影响一部分区分结果(王韫佳，覃夕航，2015)。因此，如何评估经典范畴知觉实验中区分结果的可靠性，是一个很值得研究的问题，它既有理论层面的重要性，也有技术层面的必要性。而从前人对区分实验结果不符合理论预期的分析看，在两个相似调形的范畴感知实验中，区分结果与辨认结果的矛盾几乎是传统的范畴感知实验范式所难以克服的。此外，还有一个技术层面的问题值得研究，即区分实验结果中的个体差异，刘思维(2015)在她的研究中尝试性地对个体差异进行了简单分析，她发现，区分实验结果中没有出现区分高峰，并不一定是听者不能很好地区分跨范畴的刺激，有时候是因为不同听者感知到的范畴边界的位置不同，因此在结果的均值中，个体所具有的峰值就被削平了。

1.2　ERP 研究

随着脑神经科学的发展，电生理学的研究方法使得探究声调感知行为背后的神经机制成为可能。ERP 技术具有较为精确的时间分辨率，可以用此探究声调范畴知觉的加工过程。到目前为止，已有很多研究使用 ERP 技术对声调范畴知觉模式展开探究(例如 Chandrasekaran, Krishnan, & Gandour, 2009; Kaan, Barkley, Bao, & Wayland, 2008; Shen, 2015; Wu et al., 2015; Xi et al., 2010; Zhang et al., 2011; Zhang, Xi, Wu, Shu, & Li, 2012)。

一些 ERP 研究表明，声调母语者对母语声调的知觉呈现范畴化特征，具体表现为，与范畴内偏差刺激相比，范畴间的偏差刺激可以诱发更大

的电生理反应(Shen, 2015; Xi et al., 2010; Zhang et al., 2012)。例如，在前注意阶段，Xi 等人(2010)考察了汉语普通话阳平−去声连续统的范畴知觉，结果表明，与标准刺激相比，在大脑前−中区域的双侧脑区，范畴内和范畴间刺激都可以诱发出 MMN 成分，并且，范畴间刺激在左侧脑区能够诱发更大的 MMN 成分。基于 ERP 的研究还可以区分大脑对声学信息和音系范畴的不同加工，Yu, Wang, Li 和 Li (2014)使用与 Xi 等人(2010)相类似的连续统设计，分别考察了这两种层面的信息对汉语声调范畴知觉的作用，结果表明，音系范畴和声学差异对 MMN 成分的波幅都存在显著的影响，而对于 MMN 成分的潜伏期，只有音系范畴存在主效应。他们认为，这两种层面的信息在汉语声调知觉过程中起着不同的作用，声学信息只会影响声调加工的程度，而音系范畴既会对声调加工程度产生作用，也会影响声调加工的时程。综上来看，在非注意加工阶段，声调的知觉过程不仅仅依据声学信息，也在一定程度上依赖语言层面的信息(Jia, Tsang, Huang, & Chen, 2015)。

在注意加工阶段，范畴内偏差刺激和范畴间偏差刺激也会诱发不同的电生理反应(Shen, 2015; Zhang et al., 2012)。Zhang 等人(2012)的研究结果表明，被试对汉语声调连续统进行反应时，范畴内偏差刺激在左、右双侧都诱发了 N2a 和 P3a 成分，并且范畴间偏差刺激在大脑左侧诱发了更大的电生理反应，反映了注意加工阶段对语言信息的加工。Shen (2015)也指出，不管是在早期的前注意阶段，还是较晚的注意阶段，普通话母语者对汉语声调的知觉都呈现出范畴型知觉模式，范畴间偏差刺激比范畴内偏差刺激诱发了更大的电生理反应。

然而，另外一些研究却得到了不同的结果。Chandrasekaran 等人(2009)对汉语普通话阴平和阳平(T1 和 T2)的感知研究结果表明，范畴内条件下诱发的 MMN 大于范畴间条件下的结果。该研究结果与其他结果不同，可能是因为刺激差异导致的。在大多数的研究中，范畴内和范畴间的刺激与标准刺激在声学层面的差异是相同的，即标准刺激在声学空间上居于范畴内偏差刺激和范畴间偏差刺激中间的位置。但该研究所用的范畴间刺激与标准刺激在声学上更加相似，因此得到了相反的结果。Zheng 等人(2014)对比了语言刺激与非语言刺激的范畴知觉，他们发现，非语言的范畴间刺激比范畴内刺激诱发了更大的晚期失匹配成分(late mismatch component)，而语言刺激的知觉则不存在这种范畴效应。

与行为实验的情况相类似，基于 ERP 的研究结果也出现了分歧。由上文可知，部分研究结果表明，声调的知觉具有范畴化特征，表现为范畴间偏差刺激比范畴内偏差刺激诱发的电生理反应更大(Shen, 2015; Xi et al., 2010, Zhang et al., 2012)；而另一些研究则没有得到这种范畴效应(Zheng et al., 2014)，甚至得到与之相反的结果(Chandrasekaran et al., 2009)。我们认为，造成这些差异的重要原因是声调加工过程存在阶段性，即前注意阶段和注意阶段。在注意加工阶段，声调的知觉模式是范畴型的，即在该阶段，声学信息和音系信息都得以加工，以往的研究对此争议较小。这里值得注意的是前注意阶段的结果，在前注意阶段，是否存在音系信息的加工是问题的关键所在。如果该阶段能够对音系层面信息的加工，那么范畴间刺激诱发的电生理反应就会更大；如果该阶段只对声学层面的信息进行加工，范畴效应就会较小甚至不出现。到目前为止，在前注意阶段是否存在语言层面的信息加工，学界的看法仍未达成一致。

对于语音信号的加工过程，有研究者提出了两阶段模型(two-stage model) (Luo et al., 2006)，该模型认为，在早期自动加工阶段，右脑对低层次的声学信息进行加工，因此加工优势位于大脑右半球；但到了注意阶段，较高层次的语言信息获得加工，加工优势位于大脑左半球(Luo et al., 2006)。前文所述的研究结果部分验证了该模型的合理性，但是，前注意阶段是否存在语言信息的加工，仍有待于进一步的研究。

2 影响声调范畴知觉的非声调因素

上文讨论了声调本身的特征以及声调加工的阶段性对声调范畴知觉的影响。本节将讨论影响声调知觉模式的几种非声调因素：母语背景、年龄、声调所处语境以及刺激属性(语言刺激或非语言刺激)。

2.1 母语背景对声调范畴知觉的影响

大量行为实验的结果表明，被试的语言经验会影响声调的知觉。与非声调母语者相比，声调母语者感知母语声调时的范畴化程度更高(例如，

普通话 vs 英语：Mattock & Burnham, 2006; Xu, Gandour, & Francis, 2006; 普通话 vs. 法语：Hallé et al., 2004)。而非声调语言的母语者对声调进行范畴感知时，存在一定的困难(Hallé et al., 2004; Zou, Zhang, & Cao, 2012)。Hallé 等人(2004)对中国台湾汉语母语者的研究显示，法语母语者对声调的感知主要依赖物理层面的声学参数的差异，对声调范畴的感知明显弱于中国台湾汉语母语者。Zou 等人(2012)对普通话阳平−上声连续统的研究结果也表明，汉语母语者呈现范畴知觉，而日语母语者的范畴化倾向不明显。

母语背景会对声调范畴知觉产生影响，那么，母语的声调经验是否可以迁移到非母语声调的知觉过程中呢？张林军(2010)考察了韩国、日本和泰国留学生对汉语声调的感知，结果表明，零起点的韩国和日本留学生的知觉是连续型的，泰国留学生则表现出一定的范畴化特征，但与汉语母语者的知觉模式存在差异。Peng 等人(2010)进一步比较了语言和方言背景对汉语阳平−阴平连续统和去声−阴平连续统知觉的影响，结果表明，普通话、粤语和德语三组母语者，范畴边界的位置没有差异，但声调母语者的范畴边界宽度显著小于德语母语者，这表明声调语言母语者感知这两种声调的范畴化程度比非声调母语者高，而两种方言母语者之间的范畴化程度没有差异。Zheng (2010)比较了粤语、普通话和法语母语者对粤语平调的知觉模式，结果表明，粤语母语者呈现出范畴知觉模式，普通话母语者的范畴化特征较弱，法语母语者则是连续知觉模式。Shen 和 Froud (2016)的研究考察了三组被试(汉语水平较高的英语母语者、完全不懂汉语的英语母语者和汉语母语者)对普通话声调的范畴知觉模式，结果显示，汉语母语者表现出典型的范畴知觉模式，不懂汉语的英语母语者辨认的范畴边界比其他两组被试更宽，且没有出现区分高峰。汉语水平较高的英语母语者呈现出与汉语母语者相似的范畴知觉模式。王韫佳、刘思维和卿玮(2017)对两种汉语方言研究的结果表明，重庆话母语者感知普通话阴平和阳平的模式为范畴型，但范畴化程度不及普通话母语者，她们认为这个差异是重庆话声调系统中不存在高平调和升调的对立所导致的。

语言经验或者母语背景对语音范畴知觉的作用在 ERP 研究中也被发现。大量实验结果表明，音段(辅音、元音等)和超音段(语调、声调等)加工过程中诱发的知觉电生理反应都会受到被试语言经验的影响(Bidelman, Gandour, & Krishnan, 2011; Bidelman & Lee, 2015; Chandrasekaran, Krishnan, & Gandour, 2007; Kaan et al., 2008; Krishnan, Gandour, Xu, & Suresh, 2017; Xu, Gandour, Talavage et al., 2006)。例如，Chandrasekaran 等人(2007)的研究结果表明，在 T1/T3(T3 为偏差刺激)条件下，普通话母语者比英语母语者产生了更大的 MMN 成分。Bidelman 和 Lee (2015)考察了母语背景和上下文语境对汉语阳平−上声连续统知觉的影响，结果表明，与英语母语者相比，普通话母语者在进行声调范畴判断时，产生了更大的脑电反应，因此他们认为，早期的听觉皮质活动受到母语背景的影响。Zheng 等人(2014)发现，在对非语言刺激的非注意加工过程中，粤语母语者在范畴间刺激条件下产生了更大的电生理反应，而普通话母语者则没有出现范畴效应。造成该结果的原因可能是，粤语母语者受到母语音系的影响，能够更好地区分音高高度和斜率的变化。Zheng, Minett, Peng 和 Wang (2012)考察了普通话母语者和粤语母语者对粤语声调的范畴知觉模式，结果表明，在注意条件下，范畴间偏差刺激和范畴内偏差刺激诱发的电生理反应也会受到被试语言背景的影响，只有粤语被试表现出了范畴效应，范畴间偏差刺激诱发了更大的 P300 成分。

从行为结果可以看出，母语的声调经验能够扩展到非母语声调的知觉中，声调语言和非声调语言的被试感知非母语声调存在差异，非声调母语的二语习得者能够发展出声调的范畴知觉模式。但需要注意的是，母语声调系统的复杂程度以及母语声调中的对立种类也会影响到声调的知觉模式。例如，Peng 等人的结果表明，普通话和粤语母语者对普通话阳平−阴平和去声−阴平连续统的知觉模式没有显著差异(Peng et al., 2010)，但 Zheng 的研究结果显示，普通话母语者感知粤语平调的范畴化程度比粤语母语者低(Zheng, 2010)。我们认为，该结果可能是普通话和粤语声调系统的差异造成的，粤语有 9 个声调，其中有 6 种声调是依靠音高互相区别的，而普通话只有 4 个声调。在感知声调对立种类比母语少的普通话声调时，粤语母语者与普通话母语者没有差异；而在感知声调对立种类比母语多的粤语声调时，普通话母

语者的范畴化程度就比粤语母语者低了。上文提到的王韫佳等人(2017)的研究结果表明，即便母语声调系统中的声调音位数目与非母语者声调一致(重庆话和普通话都是 4 个声调)，如果对立的具体情况不同，在感知非母语声调时范畴化程度依然会稍低。

基于 ERP 的研究结果对于母语经验在声调范畴知觉中的作用有了更深入的揭示，大量结果表明，在前注意阶段，母语背景对声调知觉的影响已经开始起作用，早期的皮质活动也会受到母语经验的影响。我们认为，这些结果可以有两种解释：一种解释是，在前注意阶段，对声调的知觉既有声学信息的加工，也有语言信息的加工，因此母语背景可以通过语言信息的加工过程对知觉模式产生影响。具体来说，声调母语者由于母语系统中存在这种音系层面的范畴差异，在声调知觉过程中更加容易感知到范畴的差异，从而产生更大的电生理反应，而非声调母语者长时记忆中没有这种范畴概念，因此也不会出现范畴效应。第二种解释是，在前注意阶段，只有声学信息的加工，没有语言信息的加工。母语经验之所以会存在显著影响，是因为声调母语者长期受到母语声调系统的影响，对某些声学参数的变化更加敏感，因此，即便不存在语言层面的信息加工，声调母语者对某些声学参数的高度敏感也会导致范畴效应。总之，在非注意阶段，母语背景是通过何种方式对声调的范畴知觉产生影响的，仍有待于更加深入的实验研究。

2.2 年龄对声调范畴知觉的影响

年龄也是影响声调范畴知觉的一个重要因素，在生长发育的不同阶段，声调的知觉会呈现出不同的模式。语音知觉发育的研究表明，在出生后的第一年中，婴儿会出现语音知觉重组(perceptual reorganisation)现象，即在出生初期，婴儿对母语和非母语的语音都能够区分，但随着母语的输入的增加，婴儿会逐渐失去辨别非母语语音的能力。以往的研究表明，婴儿对非母语辅音知觉能力的下降出现在 6 个月左右(Polka & Werker, 1994)，而对元音的知觉会更早表现出该特点(Rvachew, Alhaidary, Mattock, & Polka, 2008)。

关于声调知觉，Mattock 和 Burnham (2006)考察了 6 个月和 9 个月大的英语婴儿和汉语婴儿对泰语升调与降调、升调与平调的区分能力，结果表明，中国婴儿的声调知觉能力不断增强，而英语婴儿的声调知觉能力逐渐减弱，6~9 月龄时两者的声调识别成绩已表现出显著性差异。为了进一步验证非声调语言的婴儿对声调辨别能力的下降是否早于 6 个月，Mattock, Molnar, Polka 和 Burnham (2008)考察了 4 个月、6 个月和 9 个月的英语和法语婴儿，结果表明，英语和法语婴儿声调辨别能力的下降的确出现于 6~9 月之间。Yeung, Chen 和 Werker (2013)比较了母语为英语、普通话和粤语的婴儿对粤语声调的知觉能力，结果表明，英语婴儿在 4~9 个月期间对声调的区分能力持续下降，在此期间普通话婴儿和粤语婴儿都能够区分声调，但不同方言背景的两组婴儿在 4 个月左右就开始表现出母语偏好。Liu 和 Kager (2014)的结果有所不同，他们发现，非声调语言(荷兰语)的婴儿对声调的感知表现出 U-型的发展模式，即婴儿在 5~6 个月时可以对声调进行区分，到 9 个月左右，他们对声调的敏感有所退化，但长到 17~18 个月时，声调的知觉能力又得以恢复。

近年来，学界也开始使用电生理的方法研究婴儿言语知觉过程的神经机制。Friederici, Friedrich 和 Christophe (2007)比较了 4 个月大的德语婴儿和法语婴儿对重音模式的感知，ERP 结果表明，母语背景会影响婴儿感知重音时的大脑反应，婴儿对母语的重音模式更加敏感。Cheng 等人(2013)考察了汉语新生儿和 6 月龄婴儿感知声调时的脑电反应，结果表明，在 T1/T3(T3 为标准刺激)条件下，新生儿产生了正向的失匹配反应(positive mismatch response, P-MMR)，而 6 个月大的婴儿则产生了与成人类似的失匹配负波(MMN)；在 T2/T3(T3 为标准刺激)条件下，新生儿没有出现显著的 MMR，而 6 个月大的婴儿却出现了 P-MMR。他们认为婴儿的年龄和刺激的差异程度都会影响婴儿感知声调时的脑电反应。该研究让我们联想到前文介绍的 Chandrasekaran 等人(2007)对成人感知 T1/T3 和 T2/T3 的 ERP 研究，成人的结果表明，声学相似度低的偏差刺激能够诱发幅值更大的 MMN，而婴儿对于声学差异的反应显然更为复杂，偏差刺激诱发的 MMR 同时受到婴儿年龄和刺激声学相似度的影响。

研究婴儿的声调知觉问题，其中一个重要的问题则是，婴儿何时能够区分母语声调和非母语声调，对非母语声调的辨别能力何时开始下降。

根据上文所介绍的研究，婴儿声调知觉的发展模式与元音和辅音都有所不同，婴儿对非母语声调敏感程度的下降比元音和辅音出现得更早。这在一定程度上可以说明，不同层次的语音知觉具有不同的发展模式。非声调语言的婴儿对声调的感知呈现出 U-型发展模式，而如前文所述，成年的非声调语言与声调语言母语者对声调的知觉也具有不同的特征。因此，我们认为，婴儿在后期所谓的声调知觉能力的恢复，可能仅仅是对声学信息知觉的恢复，这种“恢复”并不意味着非声调语言的婴儿重新发展了对声调音位层面的知觉加工能力。

由于实验条件的限制，目前学界对婴儿声调知觉的 ERP 研究仍比较少。对不同的声调进行感知时，4 月龄婴儿甚至新生儿的大脑反应已经表现出差异，这一结果与行为实验的结果是一致的。关于声调语言和非声调语言的婴儿在感知声调时的神经机制，还有待进一步的研究。

2.3　声调语境(tonal context)对声调范畴知觉的影响

声调的范畴知觉不仅仅依赖音节内部本身的基频线索，上下文语境中的音高变化也起着非常重要的作用。声调又可以分为两种类型：平调和曲折调，从前人的研究结果看，上下文的声调语境对平调和曲折调的影响是不对等的，一般来说，声调语境对平调的影响要大于曲折调。例如，粤语平调的知觉依赖于上下文语境(Francis, Ciocca, Wong, Leung, & Chu, 2006; Wong & Diehl, 2003; Zhang, Peng, & Wang, 2012)。Francis 等人(2006)发现，当上下文声调的基频较低时，目标声调刺激更可能被感知为高调；而当上下文声调的基频较高时，目标声调刺激则被感知为低调的可能性更大。Wong 和 Diehl (2003)使用粤语的三种平调作为目标刺激，他们发现，同一目标刺激在基频较低的声调语境中，99.5%被感知为高平调；而在基频较高的语境中，被感知为低平调的可能性为 95.8%；当语境的平均基频位于中间位置时，则 91.9%的情况会被感知为中平调。Zhang 等人(2012)的结果也得到了类似的对比效应。声调语境也会影响粤语平调的范畴知觉模式，粤语平调的知觉在独立音节和句末条件下，呈现出非范畴化特征，而在句中条件下，则呈现出范畴知觉模式(Francis et al., 2003; Zheng, 2010)。Zheng (2010)还发现，目标声调前、后语境对知觉的作用是不对等的，目标声调后接音节基频的高低会对目标声调的知觉模式产生影响，而前接音节的作用不显著。

与平调的知觉相比，上下文语境对曲折调的作用更加复杂。一些研究表明，上下文语境对普通话声调的感知没有显著作用。Fox 和 Qi (1990)考察了普通话阴平−阳平连续统的知觉，并比较了目标刺激单独呈现和将目标刺激与前接音节同时呈现的结果，他们发现，两种条件对声调知觉的影响较小且不稳定。其他一些研究却表明，语境能够对声调知觉的范畴化程度和范畴边界产生作用。Chen 和 Peng (2016)对普通话阴平−阳平连续统的行为研究表明，目标声调的前接语境能够影响声调范畴知觉的边界位置：当前接音节基频较高时，目标音节更倾向于被感知为阳平；如果前接音节基频较低，则更倾向于被感知为阴平。Bidelman 和 Lee (2015)的 ERP 研究结果表明，母语背景和上下文语境对普通话平−上声连续统知觉都存在显著作用；语境对母语者的影响更大，具体表现为，当目标声调与前接音节同时呈现时，普通话母语者感知目标声调的范畴化程度更高。

总的说来，语境对平调和曲折调知觉的作用有所不同，我们认为有以下几种因素在起作用。首先，曲折声调的知觉，更多依赖声调本身的曲拱特征，因为不同的曲折调以及曲折调与平调之间的区分，主要是依赖曲拱特征(包括曲拱的有无)。第二，与上面的因素相关的是，平调的感知更多依赖调阶的高低，而高低的判断是相对的而非绝对的，因此语境对平调知觉的作用会更为明显。第三，声调的感知具有“延后性”，即，对于一个声调的高低判断是在后接声调出现之后才实现的，因此前接声调对于目标声调感知的作用远不如后接声调的作用大。

2.4　刺激属性(speech vs. nonspeech)对声调范畴知觉的影响

声调作为一种超音段特征，必须负载于音段之上。在以往的研究中，负载声调的刺激既有语言刺激也有非语言刺激，而前人的结果表明，刺激属性会影响声调知觉的范畴化(Xu, Gandour, & Francis, 2006)。

从行为结果来看，不管声调负载于语言刺激还是非语言刺激，声调母语者对母语声调都能较好地进行区分，并且对语言刺激条件的区分能力更强(Lee, Vakoch, & Wurm, 1996)，范畴化程度更高

(Xu, Gandour, & Francis, 2006; Zheng et al., 2012)。非声调母语者感知声调或声调母语者感知非母语声调，语言刺激与非语言刺激的差异不明显，非声调母语者对非语言刺激的知觉甚至还呈现出更强的范畴化特征。例如，Xu 等人(Xu, Gandour, & Francis, 2006)考察了普通话和英语被试对普通话阴平−阳平连续统的知觉，结果表明，不管是语言刺激还是非语言刺激，普通话母语者的感知结果都呈现出范畴化特征；但是英语母语者对非语言刺激的知觉呈现出更强的范畴化特征。

基于 ERP 的研究结果与行为实验的结果有所不同。Ren, Yang 和 Li (2009)考察了前注意阶段语言刺激和非语言刺激对音高知觉模式的影响，结果表明，不管音高负载于语言刺激还是非语言刺激，两者诱发的 MMN 成分都没有显著差异，这说明，被试在前注意阶段仅对声学信息进行加工，在之后阶段的加工中，刺激的语言功能才会起作用。然而，Zheng 等人(2014)的结果却表明，对于非语言刺激，范畴间刺激比范畴内刺激诱发了更大的晚期失匹配成分(late mismatch component)，而对于语言刺激，则不存在这种范畴效应。他们认为，与语言刺激相比，非语言刺激可以诱发更大的范畴效应，可能是由于在前注意阶段的声调知觉主要受到谐波结构的影响，非语言刺激的谐波结构更简单更有规律性，因此范畴效应更大。另一种解释是听觉皮层的激活以及失匹配成分的产生受到了电影及字幕等视觉刺激知觉的抑制——尽管在语言和非语言条件下，被试都会关注字幕，但在语言刺激条件下，字幕的干扰可能更大。另外有研究表明，在注意加工阶段，与非语言刺激相比，语言刺激的范畴效应更大(Zheng et al., 2012)。

从以往的行为研究来看，刺激属性和母语背景都会对声调的范畴知觉产生作用，两因素之间还存在交互作用。仅当声调母语者感知母语声调时，语言刺激才能够有助于声调知觉的范畴化，也就是说，只有在这种条件下，被试长时记忆中的声调原型才会被激活，并作用于声调的范畴知觉。而非声调母语者感知声调或声调母语者感知非母语声调，语言刺激和非语言刺激没有差异，这是因为被试长时记忆中不存在与感知目标相似的声调原型，无法进行匹配，被试只能依靠声学信息进行知觉判断，因此语言刺激与非语言刺激并不存在显著差异。基于 ERP 研究结果与行为结果出现了分歧，且不同研究者在 ERP 研究中也得到了不同的结果，争议的焦点在于，在非注意阶段，刺激属性是否能够作用于声调的范畴知觉。

3 学界争论与未来展望

本文对声调知觉的范畴化问题和影响声调范畴知觉的因素进行了评述，关于声调的范畴知觉，学界较为关注以及迄今为止尚未解决的问题如下：

1)决定声调加工范畴化程度的最重要因素是什么？声调范畴化加工过程的神经机制是怎样的？

2)母语背景对声调范畴知觉的影响是否受到母语声调系统复杂程度的制约？母语背景作用于声调范畴知觉的内在神经机制是怎样的？

3)婴儿声调范畴知觉能力的发展模式是怎样的？在声调知觉过程中，婴儿和成人的神经机制有哪些相似性和差异？

4)母语背景、年龄、语境和刺激属性是如何共同作用于声调范畴知觉的？

如前文所述，虽然学界对声调范畴加工的行为和神经机制研究已经取得一些成果，但仍有诸多问题存在争议，我们认为，未来的研究可以从以下几个方面进一步展开。

首先，调形的相似度与声调知觉模式之间的关系有待通过更多的实验尤其是电生理实验加以验证。目前行为实验的研究结果表明，曲拱特征不同的声调的知觉呈现出明显的范畴化特征，而曲拱特征相似度较高的声调则感知为连续型或弱范畴化型，具体表现是辨认结果呈现出范畴化特点，而区分结果无法通过辨认结果进行预测，或偏差刺激没有诱发出 MMN。如前文所指出的那样，区分实验结果受到了诸多因素的影响，区分任务本身又具有一定的复杂性，尤其是个体差异历来被人们所忽视，因此对于区分实验结果不宜像过去那样采用均值来对结果进行简单评估，而是应该从刺激排序、刺激在连续统中的位置以及个体差异等几个方面进行深入细致的分析。

如上文所述，行为实验的结果在很大程度上受到了行为实验无法克服的设计缺陷的作用，因此基于行为实验结果提出的假设是否正确，还需通过实验技术的更新加以验证。曲拱特征的相似度，即声学层面的相似度，是否能够在大脑加工过程中反映出来，影响声调范畴知觉的因素到底是什么，都有待借助电生理技术进行更深入的研

究。而从目前的研究现状来看, 这方面的研究几乎还是一个空白。

第二, 关于母语背景对声调范畴知觉的影响, 未来研究可以关注一下几个问题：第一, 以往研究主要关注声调语言和非声调语言被试之间的差异, 我们认为这是远远不够的。未来的研究中需要充分考虑不同声调语言声调系统的复杂程度。比如, 母语声调系统的复杂程度是否会影响被试感知声调的能力？是否母语声调系统越复杂, 在感知非母语声调时就会越敏感？第二, 母语背景对声调范畴知觉的影响, 主要发生在认知加工的哪一层面？哪一阶段？母语经验是否只会影响语言信息的加工？是否也会影响到声学参数的感知？只有把这些具体问题研究清楚了, 才能更好地解释母语背景对声调范畴化知觉的作用。另外, 方言背景对声调感知的作用, 学界研究仍比较少。汉语的方言资源非常丰富, 而各汉语方言在声调种类和调形上千差万别。未来的研究应该加大力度研究汉语方言母语者对非母语方言声调的知觉, 以期在声调知觉的细节上得到更深入和更全面的结果。

第三, 关于年龄对声调知觉能力的影响, 如前文所述, 有研究认为婴儿的声调知觉能力呈 U-型发展模式。我们认为这只是一个初步的结论, 有待加以精准化研究, 并与成人非母语声调的感知研究进行联系。所谓精准化研究, 是对一系列细节问题的深入讨论, 例如, 声调感知能力的 U-型发展与整个语音能力的发展之间是什么关系, 婴儿非母语声调感知能力的恢复期何时结束、在何种水平上结束, 婴儿对非母语声调的感知特征是什么等等。另外, 声调语言和非声调语言的婴儿对声调感知过程中的神经机制, 也有待进一步研究。这里最重要的问题是, 声调语言的婴儿对母语声调的感知具有相当的复杂性, 或者说婴儿和成人对母语声调的感知具有不同的特点, 这些不同特点都有哪些, 它们是如何逐渐消失的, 都需通过更多的电生理实验加以深入研究。

第四, 关于语境对声调感知范畴化的影响, 目前学界得到的结果是一致的, 即, 在孤立条件或者句末、词末条件下未呈现范畴化知觉模式的声调, 在句中或者词首位置的感知都会出现范畴化程度的加强。如上文所述, 我们猜测这与声调感知的延后性相关。但“延后性”只能对前后语境影响的不对称给予解释, 却无法说明知觉模式从非范畴型到范畴型的改变。这一问题的解决除了需要在实验设计上有所突破外, 纯理论层面的讨论也许是更加重要的。

最后, 母语背景、年龄、语境和刺激属性对声调范畴知觉的作用是相互影响的, 比如, 刺激类型对范畴知觉的影响在声调母语者和非声调母语者之间存在差异。那么, 这四种因素是如何单独起作用, 它们之间又是如何共同起作用的呢？还有有待于进一步的研究。

参考文献

蔡雯清. (2016). *合肥话单字调的知觉研究*(硕士学位论文). 北京大学.

刘思维. (2015). *汉语声调相异调形和相似调形声调的范畴感知研究*(博士学位论文). 北京大学.

覃夕航. (2012). *母语经验对汉语普通话声调范畴化感知的影响*(硕士学位论文). 北京大学.

王韫佳, 李美京. (2010). 调型和调阶对阳平和上声知觉的作用. *心理学报, 42*, 899–908.

王韫佳, 刘思维, 卿玮. (2017). 从范畴感知看重庆话阴平和阳平的调型——兼论重庆话阳平和上声演变的动因. 见 李爱军(编). *中国语音学报(第 7 辑)* (pp. 18–27). 北京: 中国社会科学出版社.

王韫佳, 覃夕航. (2015). 普通话单字调阳平和上声的辨认及区分——兼论实验设计对声调范畴感知结果的影响. *语言科学, 14*, 337–352.

张林军. (2010). 日本留学生汉语声调的范畴化知觉. *语言教学与研究,* (3), 9–15.

Abramson, A. S. (1979). The noncategorical perception of tone categories in Thai. In B. Lindblom & S. Öhman (Eds.), *Frontiers of speech communication research* (pp. 127–134). London, UK: Academic Press.

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience, 23*, 425–434.

Bidelman, G. M., & Lee, C. C. (2015). Effects of language experience and stimulus context on the neural organization and categorical perception of speech. *NeuroImage, 120*, 191–200.

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007). Mismatch negativity to pitch contours is influenced by language experience. *Brain Research, 1128*, 148–156.

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain and Language, 108*, 1–9.

Chen, F., & Peng, G. (2016). Context effect in the categorical perception of mandarin tones. *Journal of Signal Processing Systems, 82*, 253–261.

Cheng, Y. Y., Wu, H. C., Tzeng, Y., Yang, M. T., Zhao, L. L., & Lee, C. Y. (2013). The development of mismatch responses to mandarin lexical tones in early infancy. *Developmental Neuropsychology, 38*, 281–300.

Fox, R. A., & Qi, Y. Y. (1990). Context effects in the perception of lexical tone. *Journal of Chinese Linguistics, 18*, 261–284.

Francis, A. L., & Ciocca, V. (2003). Stimulus presentation order and the perception of lexical tones in Cantonese. *The Journal of the Acoustical Society of America, 114*, 1611–1621.

Francis, A. L., Ciocca, V., & Ng, B. K. C. (2003). On the (non)categorical perception of lexical tones. *Perception & Psychophysics, 65*, 1029–1044.

Francis, A. L., Ciocca, V., Wong, N. K. Y., Leung, W. H. Y., & Chu, P. C. Y. (2006). Extrinsic context affects perceptual normalization of lexical tone. *The Journal of the Acoustical Society of America, 119*, 1712–1726.

Friederici, A. D., Friedrich, M., & Christophe, A. (2007). Brain responses in 4-month-old infants are already language specific. *Current Biology, 17*, 1208–1211.

Fry, D. B., Abramson, A. S., Eimas, P. D., & Liberman, A. M. (1962). The identification and discrimination of synthetic vowels. *Language and Speech, 5*, 171–189.

Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics, 32*, 395–421.

Jia, S., Tsang, Y. K., Huang, J., & Chen, H. C. (2015). Processing Cantonese lexical tones: Evidence from oddball paradigms. *Neuroscience, 305*, 351–360.

Kaan, E., Barkley, C. M., Bao, M. Z., & Wayland, R. (2008). Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: An event-related potentials training study. *BMC Neuroscience, 9*, 53.

Krishnan, A., Gandour, J. T., Xu, Y., & Suresh, C. H. (2017). Language-dependent changes in pitch-relevant neural activity in the auditory cortex reflect differential weighting of temporal attributes of pitch contours. *Journal of Neurolinguistics, 41*, 38–49.

Lee, Y. S., Vakoch, D. A., & Wurm, L. H. (1996). Tone perception in Cantonese and mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research, 25*, 527–542.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*, 358–368.

Liu, L. Q., & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition, 133*, 385–394.

Luo, H., Ni, J. T., Li, Z. H., Li, X. O., Zhang, D. R., Zeng, F. G., & Chen, L. (2006). Opposite patterns of hemisphere dominance for early auditory processing of lexical tones and consonants. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 19558–19563.

Mattock, K., & Burnham, D. (2006). Chinese and English infants′ tone perception: Evidence for perceptual reorganization. *Infancy, 10*, 241–265.

Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition, 106*, 1367–1381.

Peng, G., Zheng, H. Y., Gong, T., Yang, R. X., Kong, J. P., & Wang, W. S. Y. (2010). The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics, 38*, 616–624.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 421–435.

Ren, G. Q., Yang, Y., & Li, X. (2009). Early cortical processing of linguistic pitch patterns as revealed by the mismatch negativity. *Neuroscience, 162*, 87–95.

Rvachew, S., Alhaidary, A., Mattock, K., & Polka, L. (2008). Emergence of the corner vowels in the babble produced by infants exposed to Canadian English or Canadian French. *Journal of Phonetics, 36*, 564–577.

Shen, G. N. (2015). *Perceptual learning of lexical tone categories: An ERP study* (Unpublished doctorial dissertation). Columbia University.

Shen, G. N., & Froud, K. (2016). Categorical perception of lexical tones by English learners of mandarin Chinese. *The Journal of the Acoustical Society of America, 140*, 4396–4403.

Studdert-Kennedy, M., & Shankweiler, D. (1970). Hemispheric specialization for speech perception. *The Journal of the Acoustical Society of America, 48*, 579– 594.

Wang, W. S. Y. (1976). Language change. *Annals of the New York Academy of Sciences, 280*, 61–72.

Wong, P. C. M., & Diehl, R. L. (2003). Perceptual normalization for inter- and intratalker variation in Cantonese level tones. *Journal of Speech, Language, and Hearing Research, 46*, 413–421.

Wu, H., Ma, X. H., Zhang, L. J., Liu, Y. Y., Zhang, Y., & Shu, H. (2015). Musical experience modulates categorical perception of lexical tones in native Chinese speakers.

Frontiers in Psychology, 6, 436.

Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience, 170*, 223–231.

Xu, Y. S., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *The Journal of the Acoustical Society of America, 120*, 1063–1074.

Xu, Y. S., Gandour, J. T., Talavage, T., Wong, D., Dzemidzic, M., Tong, Y. X., … Lowe, M. (2006). Activation of the left planum temporale in pitch processing is shaped by language experience. *Human Brain Mapping, 27*, 173–183.

Yeung, H. H., Chen, K. H., & Werker, J. F. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language, 68*, 123–139.

Yu, K. K., Wang, R. M., Li, L., & Li, P. (2014). Processing of acoustic and phonological information of lexical tones in mandarin Chinese revealed by mismatch negativity. *Frontiers in Human Neuroscience, 8*, 729.

Zhang, C. C., Peng, G., & Wang, W. S. Y. (2012). Unequal effects of speech and nonspeech contexts on the perceptual normalization of Cantonese level tones. *The Journal of the Acoustical Society of America, 132*, 1088–1099.

Zhang, L. J., Xi, J., Xu, G. Q., Shu, H., Wang, X. Y., & Li, P. (2011). Cortical dynamics of acoustic and phonological processing in speech perception. *PLoS One, 6*, e20963.

Zhang, L. J., Xi, J., Wu, H., Shu, H., & Li, P. (2012). Electrophysiological evidence of categorical perception of Chinese lexical tones in attentive condition. *Neuroreport, 23*, 35–39.

Zheng, H. Y. (2010). *Categorical perception of lexical tones: Behavioral and psychophysiological study* (Unpublished doctorial dissertation). The Chinese University of Hong Kong.

Zheng, H. Y., Minett, J. W., Peng, G., & Wang, W. S. Y. (2012). The impact of tone systems on the categorical perception of lexical tones: An event-related potentials study. *Language and Cognitive Processes, 27*, 184–209.

Zheng, H. Y., Peng, G., Chen, J. Y., Zhang, C. C., Minett, J. W., & Wang, W. S. Y. (2014). The influence of tone inventory on ERP without focal attention: A cross-language study. *Computational and Mathematical Methods in Medicine, 2014*, 961563.

Zou, T., Zhang, J. S., & Cao, W. (2012). A comparative study of perception of tone 2 and tone 3 in Mandarin by native speakers and Japanese learners. In *Proceedings of the 8th international symposium on Chinese spoken language processing* (pp. 431–435). Kowloon, Hong Kong, China: IEEE.

Categorical perception of lexical tone and the neural mechanisms

WU Qian[1]; WANG Yunjia[1,2]

([1] *Department of Chinese Language and Literature, Peking University, Beijing* 100871, *China*)

([2] *Center for Chinese Linguistics, Peking University, Beijing* 100871, *China*)

Abstract: Categorical perception (CP) refers to the ability that listeners can distinguish cross-category stimuli, but cannot discriminate different stimuli within the same category. The degree of categorization for lexical tone perception may be influenced by the acoustical similarity of different tones. The perception of two similar tones seems to be less categorical than that of dissimilar tones. In addition to the internal features of tones, the factors influencing the categorical perception of lexical tone include language experience, age, tonal context and stimulus type (speech or nonspeech). The use of electrophysiological techniques has advanced the study of categorical perception of lexical tone, and is expected to be helpful in solving more theoretical problems.

Key words: tone; categorical perception; neural mechanisms